Operationalization of Clinical Practice Guidelines Using Fuzzy Logic

James C.S. Liu, MD, Richard N. Shiffman, MD, MCIS
Center for Medical Informatics
Yale University School of Medicine
New Haven, Connecticut

There are a number of obstacles to successful operationalization of clinical practice guidelines, including the difficulty in accurately representing a statement's decidability or an action's executability. Both require reasoning with incomplete and imprecise information, and we present one means of processing such information. We begin with a brief overview of fuzzy set theory, in which elements can have partial memberships in multiple sets. With fuzzy inferencing, these sets can be combined to create multiple conclusions, each with varying degrees of truth. We demonstrate a fuzzy model developed from a published clinical practice guideline on the management of first simple febrile seizures. Although the creation of fuzzy sets can be an arbitrary process, we believe that fuzzy inferencing is an effective tool for the expression of guideline recommendations, and that it can be useful for the management of imprecision and uncertainty.

INTRODUCTION

There has been increasing interest in incorporating clinical practice guidelines into computer-based patient records [1, 2]. However, a number of problems have emerged from the process of operationalizing a guideline's recommendations into decision support tools [3]. One is the issue of *decidability*, or determining the exact circumstances under which an action is recommended. A vaguely phrased decision criterion such as "low-grade fever" can be difficult to interpret. Even a clearly defined sign can be more or less convincing in different patients (e.g. a "textbook" Brudzinski sign vs. a borderline or less obvious case). Another issue is the *executability* of an action, or exactly what to do if specific decision criteria are satisfied. How does one operationalize a phrase like, "IF wheezing is severe AND beta-agonists are not helpful THEN *consider* transfer to an emergency room?" Likewise, how does one differentiate *consider* from *strongly consider* or *recommend*?

In general, these issues emphasize the fact that a conventional guideline relies on deterministic, all-or-none reasoning, while clinical practice often requires reasoning with incomplete and imprecise information. The capacity for approximate reasoning is a critical component of any computer-based medical inferencing system. We describe a useful means of approximate reasoning for operationalizing clinical guideline logic.

Fuzzy Sets and Fuzzy Logic

Fuzzy sets were first described by Zadeh in 1965 [4]. He extended Aristotelian two-valued logic and conventional set theory by proposing sets with indistinct boundaries. An element in a fuzzy set has a partial membership in it, rather than all-or-none membership as in a conventional set. This degree of membership is described by a membership function, whose value lies in the continuous interval of [0, 1]. Thus in the set of tall people, Michael, who is 6'2", could have a membership of 0.8, implying that he is 80% tall. A person who stands 7' would have a higher membership value; someone 5'10" would have a lower value. This is distinct from probabilistic reasoning, in that the latter describes a likelihood of absolute membership or non-membership (Michael has an 80% chance of being absolutely tall, a 20% chance of being short), while a fuzzy membership implies that an element has an intrinsic, partial membership within the set (Michael is 80% tall). In addition, a fuzzy element can belong to multiple sets, but the sum total of the membership values in the sets that contain it can be less than or greater than 1.0 [5].

Zadeh also described a series of logical operations that can be performed on fuzzy sets. The classical NOT would be implemented by taking 1.0 minus the membership value (e.g. if Michael is 0.8 tall, he is 0.2 NOT-tall). The classical OR (or the UNION, in set theory or modified ADDITION, in probabilistic reasoning) is used to collect the elements that are in either set, and is implemented in fuzzy sets by taking the maximum membership between the two values (MAXIMA). So if Michael is 0.8 tall and 0.4 strong, he is 0.8 tall OR strong. The classical AND (set theory INTERSECTION or probabilistic MULTIPLICATION) is an operation which determines the elements which are common to two sets, and is often implemented as the minimum membership value between two or more elements (MINIMA). (In our above example, Michael would

1091—8280/97/$5.00 © 1997 AMIA, Inc.

be 0.4 tall AND strong). Using these definitions, one can derive all of the fundamental properties of classical logic and set theory, except for the Law of Contradiction (A AND NOT-A = null set) and the Law of the Excluded Middle (A OR NOT-A = universal set) [6]. After all, in our fuzzy example, Michael can be 0.2 tall AND NOT-tall, and 0.8 tall OR NOT-tall.

In a subsequent paper, Zadeh described methods of inferencing using fuzzy sets [7]. One can group fuzzy sets using MAXIMA or MINIMA operations, or create relationships between fuzzy variables with conditional IF-THEN statements. Zadeh also described mathematical operators which he called "hedges," which approximate linguistic modifiers of words (e.g. "more or less," "very," etc.). Based on these principles, Zadeh developed an elaborate system of reasoning, in which partial membership in multiple sets can result in multiple possible conclusions, each with varying degrees of truth. He also described fuzzy algorithms which extend these principles to handle more complex problems.

Zadeh's fuzzy inferencing methods have found their widest use in control systems. Fuzzy designers have taken advantage of the ability of fuzzy sets to express vague linguistic terms, and perform inferencing using expert-derived, intuitively phrased rules. They have exploited the capacity of fuzzy inferencing to create systems with a high tolerance for uncertain or incomplete information [8]. Fuzzy inferencing has principally been used in medicine in diagnosis and classification engines [9, 10, 11], in control systems [12, 13], and in pattern recognition and image enhancement [14, 15].

THE GUIDELINE

The Guidelines Review Group at the Yale Center for Medical Informatics has been evaluating the logical integrity of guidelines published by the American Academy of Pediatrics (AAP). In our evaluations, we identify decision variables and recommended actions from the guidelines, and assess their decidability and executability.

One guideline that we have analyzed concerns the AAP's recommendations for management of the first simple febrile seizure in otherwise healthy children aged 6 to 60 months [16]. The guideline (1) describes the children who are eligible for its recommendations, (2) states that routine diagnostic use of EEG's, neuroimaging, and bloodwork is not generally indicated, and (3) lists factors involved in the decision to perform lumbar puncture (LP) as part of the diagnostic evaluation.

Problems with Implementation

A number of problems arose in considering how the guideline could be implemented in a computer program. For example, the criteria for deciding whether to perform LP, and the strength of recommendation for LP, are not clear-cut. The guideline states:

> The American Academy of Pediatrics recommends, on the basis of published evidence and consensus, that after the first seizures with fever in infants *younger than 12 months*, performance of a lumbar puncture be *strongly considered,* because the clinical signs and symptoms associated with meningitis may be minimal or absent in this age group. In a child *between 12 and 18 months* of age, a lumbar puncture should be *considered,* because clinical signs and symptoms of meningitis may be subtle. In a child *older than 18 months,* although a lumbar puncture is *not routinely warranted,* it is *recommended* in the *presence of meningeal signs and symptoms* (i.e., neck stiffness and Kernig and Brudzinski signs), which are usually present with meningitis, or for any child whose history or examination result *suggests the presence of intracranial infection.* In infants and children who have had febrile seizures and have *received prior antibiotic treatment,* clinicians should be aware that treatment can mask the signs and symptoms of meningitis. As such, a lumbar puncture should be *strongly considered* (italics added) [16].

In analyzing this paragraph, we isolated three principal decision variables that factor into the appropriateness of lumbar puncture. Fuzzy inferencing can help to quantify some of the issues of decidability around each of these variables. The first fuzzy variable is the presence of *evidence suggestive of intracranial infection.* This can be open to interpretation; Brudzinski sign can be obvious in one case and merely suspicious in another. Moreover, there are less specific signs such as neck stiffness, lethargy, elevated peripheral white blood cell (WBC) count, and petechiae. The guideline does not distinguish these from more persuasive signs.

A *history of recent antibiotics* is the second fuzzy decision variable. The specific antibiotic given may not be critical, since most antibiotics commonly prescribed for children have some degree of activity against the organisms most frequently responsible for meningitis. However, the interval between the last dose of an antibiotic and the seizure is of paramount importance; a drug given yesterday is much more confounding than one given three days ago, and either is far more concerning than one given two or three weeks ago.

Our third fuzzy decision variable is the *child's age*; the guideline draws crisp distinctions between children 6-12 months old, 12-18 months old, and 18-60 months old. This reflects the observation that a child's ability to display recognizable signs of meningitis increases with time. However, this improvement occurs continuously, and not in discrete jumps. For example, there are obvious changes between a child aged 9, 15, and 36 months. On the other hand, a child of 13 months would probably act more like a child of 11 months than one of 18, though the hard-and-fast categories of the guideline do not reflect this concept.

Finally, the executability of the *recommendation for an LP* is described in terms of four fuzzy classes; an LP can be *recommended*, *strongly considered*, *considered*, or *not routinely warranted*. Again, most clinicians would think more in terms of a continuum of possibilities rather than four discrete classes, and have some intuitive means of making a decision based on combining the three input variables. The guideline offers no means of combining these variables. We believe that these decision variables and the executability of the decision for an LP are all eminently suitable for "fuzzification."

Using the above decision criteria, we reduced the paragraph from the guideline to the following five rules [details in 17]:

- IF meningeal signs are *present*, THEN LP is *recommended*.
- IF antibiotics have been given *recently*, THEN LP should be *strongly considered*.

If neither of these is true:

- IF the child is 6-12 months old, THEN LP should be *strongly considered*.
- IF the child is 12-18 months old, THEN LP should be *considered*.
- IF the child is 18-60 months old, THEN LP is *not routinely warranted*.

Applying Fuzzy Sets to the Decision Rules

Our output variable is the strength of recommendation for lumbar puncture. We will implement the decision by dividing appropriateness of LP into four fuzzy classes: (a) *not routinely warranted* (we arbitrarily assign membership values of 0 to 0.25 in this class); (b) *considered* (interval from 0.25 to 0.5); (c) *strongly considered* (interval from 0.5 to 0.75); and (d) *recommended* (interval from 0.75 to 1.0).

The three fuzzy decision variables can then be classified in terms of the degree that each of them contributes to the LP. The first variable, the *evidence for intracranial infection*, comprises consideration of its various signs, such as Kernig or Brudzinski sign, WBC count, lethargy, etc. Each of these signs can be assigned its own membership function. But not all signs are equally concerning; a clear-cut Brudzinski sign would be sufficient to recommend LP, but other signs such as a high WBC count might not raise as much concern; thus each membership function may not have the same maximum value. For example, if the maximum concern from a high WBC count were to *strongly consider* LP, or 0.75, and the leucocyte count was 11,000/µl—a 60% convincing high WBC count—the total WBC score would be 60% of the maximum 0.75, or 0.45.

Since a single convincing sign would be sufficient grounds to perform LP, we will combine the presence of these signs using a fuzzy OR (MAX) operator, where Brudzinski sign is assigned a maximum possible appropriateness of 1.0 (*recommended*), and other signs can be assigned a less concerning maximum score (e.g., 0.5). Thus if Brudzinski sign is 75% convincing and the WBC count is 60% elevated (score=0.45), the level of recommendation for an LP is the maximum of these memberships, or 0.75 (*strongly consider*).

The *interval since the patient last received antibiotics* is the second fuzzy decision variable. In the guideline, recent antibiotics can, at most, cause the clinician to *strongly consider* LP; presumably concern dwindles with increasing time since the last dose. We can therefore implement this variable with a simple straight-line membership function (Figure 1) whose value is 0.75 at time 0 and steadily decreases with time since the last dose of antibiotic.

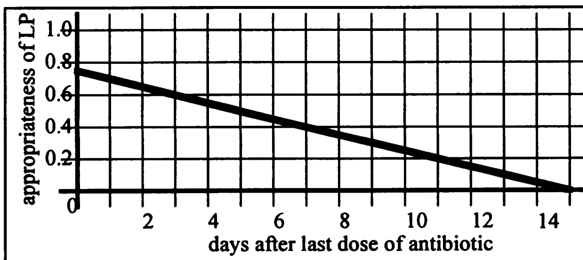

Figure 1: Interval since antibiotic vs. appropriateness of LP

The *patient's age* is the third input variable. We propose that an index of concern that continuously decreases over time would be more realistic than three large age classes. When age alone is considered as a trigger for LP, concern never rises higher than to *strongly consider* LP for a 6-month-old child; at 12 months, the physician is instructed to *consider* LP, and at 18 months, LP is *not routinely warranted.* We can therefore define a membership function that starts with a value of 0.75 at 6 months, decreases steadily to 0.5 at 12 months, and reaches 0.25 at 18 months. After that, concern steadily decreases such that it becomes minimal by the age of 60 months (Figure 2).

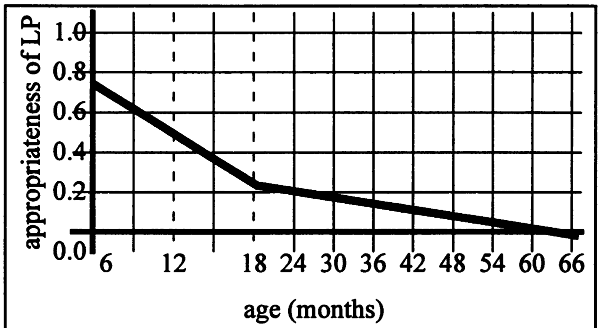

Figure 2: Patient's age vs. appropriateness of LP

Decision Making with These Fuzzy Sets

We assume that any of these three variables alone would be sufficient to trigger the LP decision in a child. This suggests that an inference engine could simply take the highest value among the three variables (fuzzy OR) to create an output score. This score can then be back-translated into a linguistic level of appropriateness for an LP.

Recommendation for LP = MAX (suspicion of intracranial infection, age, time since last antibiotic)

For example, if the patient was 17 months old, last received antibiotics 5 days ago for an episode of otitis media, and had an 80% convincing Brudzinski sign, then (1) the age score is about 0.3, (2) the antibiotic score is about 0.6, and (3) the evidence of intracranial infection score is 0.8. The output appropriateness of an LP, therefore, would be 0.8, which translates into an LP being *recommended.* However, this score can easily be changed, e.g. if the meningeal signs were less convincing, the score would decrease markedly, while if antibiotics were given more recently, the score could increase as high as 0.75, etc.

DISCUSSION

Several advantages become evident with the use of fuzzy modeling. For example:

- We believe that these continuous fuzzy variables better fit a human's approach to reasoning than the guideline's crisp categories. Decision variables gradually change from one state to another, rather than abruptly changing at arbitrary and arguable cutoff values.
- Our inferencing method uses these variables to adapt complex, nonlinear concepts—which are not as amenable for mathematical modeling—into linguistic constructs that are easy to develop and understand.
- The system does not require an assumption that our input variables are independent, as Bayesian probabilistic approaches do.
- In general, a well-designed fuzzy system tends to be more tolerant of ambiguity and uncertainty than probabilistic and conventional algorithmic approaches.
- In more complex multivariable systems, the use of fuzzy classes allows multiple options to be activated with varying strengths. The system designer can then choose a threshold membership value to determine how many options should be presented to the user, and the comparative strength of each option.

At the same time, a number of disadvantages become readily apparent with fuzzy inferencing:

- The membership functions and classes used to describe fuzzy sets are inherently arbitrary. There are multiple ways that we could assign consideration for LP (*not routinely performed* could have been assigned a value of 0, *recommended* a value of 1, and *considered* and *strongly considered* could split the range in between); those reassigned choices could then influence the design of other classes.
- Once the rules are defined, they require extensive testing to confirm their ability to describe reality accurately.
- As the number of system variables increases, the rule base explodes at an exponential rate, reducing system comprehensibility.

A number of additional tools have been described to help handle the additional complexity of more ambitious systems. For example, fuzzy decision tables can be used to help systematically analyze rules to insure against incompleteness, redundancy, or contradiction [18, 19].

We have described a method for operationalizing a current clinical practice guideline by using fuzzy logic to express the uncertainty inherent in the system. We believe that fuzzy sets make it possible to quantify a series of vague linguistic concepts. These concepts can then be combined with the flexibility of certainty factors, with more expressiveness than Dempster-Shafer theory, and fewer restrictions on variable independence than in Bayesian probability. Fuzzy inferencing thus provides another useful tool to help computerized guidelines to draw conclusions from imprecise information.

Acknowledgments

Our work was supported in part by grants 1-R55-LM05552-01A1 and T-15-LM07056 from the National Library of Medicine. Dr. Shiffman is a Robert Wood Johnson Generalist Physician Faculty Scholar.

References

1. Tierney WM, Overhage JM, Takesue BY, et al. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *J. Am. Med. Inform. Assoc.*, 1995; 2(5): 316-322.
2. Shiffman RN, Greenes RA. Improving clinical guidelines with logic and decision-table techniques: application to hepatitis immunization recommendations. *Med. Decis. Making*, 1994; 14: 245-254.
3. Horn SD. National guidelines and local action: priority-setting for the development of research-based protocols at Intermountain Health Care. In: Field MJ, ed., *Setting Priorities for Clinical Practice Guidelines*. Washington, DC: National Academy Press, 1995.
4. Zadeh LA. Fuzzy sets. *Information and Control*, 1965; 8(3): 338-353.
5. Bezdek JC. Fuzzy models: what are they, and why? *IEEE Transactions on Fuzzy Systems*, 1993; 1(1): 1-6.
6 Klir GJ, Yuan B. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice Hall Inc., 1995.
7. Zadeh LA. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973; 3(1): 28-44.
8. Munakata T, Jani Y. Fuzzy systems: an overview. *Communications of the ACM*, 1994; 37(3): 69-76.
9. Adlassnig KP, Kolarz G, Scheithauer W, et al. CADIAG: Approaches to computer-assisted medical diagnosis. *Comput. Biol. Med.*, 1985; 15(5): 315-335.
10. Binaghi E, De Giorgi O, Maggi G, Motta T, Rampini A. Computer-assisted diagnosis of postmenopausal osteoporosis using a fuzzy expert system shell. *Comput. Biomed. Res.*, 1993; 26: 498-516.
11. Fathi-Torbaghan M, Meyer D. MEDUSA: A fuzzy expert system for medical diagnosis of acute abdominal pain. *Meth. Inf. Med.*, 1994; 33(5): 522-529.
12. Ying H, McEachern M, Eddleman DW, Sheppard LC. Fuzzy control of mean arterial pressure in postsurgical patients with sodium nitroprusside infusion. *IEEE Trans. Biomed. Eng.*, 1992; 39(10): 1060-1070.
13. Schäublin J, Derighetti M, Feigenwinter P, Petersen-Felix S, Zbinden AM. Fuzzy logic control of mechanical ventilation during anaesthesia. *Br. J. Anesth.*, 1996; 77(5): 636-641.
14. Peters RM, Shanies SA, Peters JC. Fuzzy cluster analysis of positive stress tests, a new method of combining exercise test variables to predict extent of coronary artery disease. *Am. J. Cardiol.*, 1995; 76(10): 648-651.
15. Shiomi S, Kuroki T, Jomura H, et al. Diagnosis of chronic liver disease from liver scintiscans by fuzzy reasoning. *J. Nucl. Med.*, 1995; 36(4): 593-598.
16. Provisional Committee on Quality Improvement. Practice parameter: The neurodiagnostic evaluation of the child with a first simple febrile seizure. *Pediatrics*, 1996; 97(5): 769-775.
17. Shiffman, RS. Representation of clinical practice guideline knowledge in conventional and augmented decision tables. *J. Am. Med. Inform. Assoc.*, accepted for publication 1997.
18. Francioni JM, Kandel A. A software engineering tool for expert system design. *IEEE Expert*, 1988; 3(1): 35-41.
19. Vanthienen J, Wets G, Chen GQ. Incorporating fuzziness in the classical decision table formalism. *International Journal of Intelligent Systems*, 1996; 11(11): 879-891.